

UNCLASSIFIED

AD NUMBER

**AD907764**

NEW LIMITATION CHANGE

TO

**Approved for public release, distribution  
unlimited**

FROM

**Distribution authorized to U.S. Gov't.  
agencies only; Test and Evaluation; 29 JAN  
1973. Other requests shall be referred to  
Office of Naval Research, Arlington, VA.**

AUTHORITY

**ONR ltr, 29 Aug 1973**

THIS PAGE IS UNCLASSIFIED

# SELECTED STATISTICAL TECHNIQUES APPLICABLE TO ASW EXERCISE DESIGN AND ANALYSIS

AD 902264

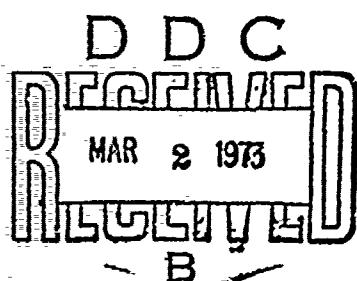
Prepared by:

Barbara W. Perneski  
Peter H. Vanderwaart  
Lyman L. McDonald

ANALYSIS & TECHNOLOGY, INC.

28 FEBRUARY 1973

PREPARED FOR:  
NAVAL ANALYSIS PROGRAMS  
OFFICE OF NAVAL RESEARCH  
ARLINGTON, VIRGINIA 22217  
CONTRACT N00014-72-C-0238  
TASK NR 364-065



"Reproduction in whole or in part is permitted  
for any purpose of the United States Government."

"Distribution limited to U. S. Government Agencies  
only; Test and Evaluation: 29 January 1973.  
Other requests for this document must be referred  
to Office of Naval Research (Code 462)."



ANALYSIS & TECHNOLOGY, INC. REPORT

Number P-66-1-72

SELECTED STATISTICAL TECHNIQUES  
APPLICABLE TO  
ASW EXERCISE DESIGN AND ANALYSIS

"Reproduction in whole or in part is permitted  
for any purpose of the United States Government."

"Distribution limited to U. S. Government Agencies  
only; Test and Evaluation: 29 January 1973.  
Other requests for this document must be referred  
to Office of Naval Research (Code 462)."

PREPARED FOR:

NAVAL ANALYSIS PROGRAMS  
OFFICE OF NAVAL RESEARCH  
ARLINGTON, VIRGINIA 22217

**UNCLASSIFIED**

Security Classification

**DOCUMENT CONTROL DATA - R & D***(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)*

1. ORIGINATING ACTIVITY (Corporate author) Analysis & Technology, Inc. P.O. Box 220 (Technology Park) North Stonington, Conn. 06359		2a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED
		2b. GROUP N/A
3. REPORT TITLE SELECTED STATISTICAL TECHNIQUES APPLICABLE TO ASW EXERCISE DESIGN AND ANALYSIS		
4. DESCRIPTIVE NOTES (Type of report and Inclusive dates) Final Report		
5. AUTHOR(S) (First name, middle initial, last name) Barbara W. Perneski Peter H. Vanderwaart Lyman L. McDonald		
6. REPORT DATE 28 February 1973	7a. TOTAL NO. OF PAGES 53	7b. NO. OF REFS 13
8a. CONTRACT OR GRANT NO. N00014-72-C-0238	9a. ORIGINATOR'S REPORT NUMBER(S) P-66-1-72	
b. PROJECT NO. RF 018-96-01	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report) N/A	
c. NR 364-065 d.		
10. DISTRIBUTION STATEMENT "Distribution limited to U.S. Government Agencies only; Test and Evaluation: 29 January 1973. Other requests for this document must be referred to Office of Naval Research (Code 462)."		
11. SUPPLEMENTARY NOTES N/A	12. SPONSORING MILITARY ACTIVITY Naval Analysis Programs (Code 462) Office of Naval Research Arlington, VA 22217	
13. ABSTRACT This report contains:  <ul style="list-style-type: none"> <li>• A method for calculating symmetric confidence intervals for search rate and mean time-to-detection for the active and passive area search missions.</li> <li>• A method for calculating approximate confidence intervals on cumulative detection probability as a function of range for the general case containing "turn-arounds" (CPAs) and "late-starters."</li> <li>• The application (and modification) of various techniques for calculating approximate confidence intervals on Mission Measures of Effectiveness that are in the form of products of proportions.</li> <li>• A discussion of the potential bias due to "late-starters" in the development of cumulative detection probability as a function of range. Recommendations for eliminating (or at least minimizing) this bias are included in this report. (U).</li> </ul>		

DD FORM 1 NOV 68 1473

REPLACES DD FORM 1473, 1 JAN 64, WHICH IS  
OBsolete FOR ARMY USE.**UNCLASSIFIED**

Security Classification

~~UNCLASSIFIED~~

Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Cumulative Detection Probability as a Function of Range  Confidence Intervals  Search Rate  Mean Time-to-Detection  Area Search  Confidence Intervals on MOEs  Sample Size Estimation						

~~UNCLASSIFIED~~

Security Classification

## SUMMARY

Several statistical problems important to submarine ASW exercise design and analysis were investigated under the research contract for updating the current SUBMARINE ANALYSIS NOTEBOOK. The application (but not the development and theory) of the research will be contained in the revised edition of the Notebook.

The underlying theory of certain results of this research needs to be published separately, since it contains new techniques developed under this contract or includes the modification and/or application of methodologies not available in standard textbooks. This paper contains this theory.

The results discussed herein include:

- A method for calculating symmetric confidence intervals for search rate and mean time-to-detection for the active and passive area search missions.
- A method for calculating approximate confidence intervals on cumulative detection probability as a function of range for the general case containing "turn-arounds" (CPAs) and "late-starters."
- The application (and modification) of various techniques for calculating approximate confidence intervals on Mission Measures of Effectiveness that are in the form of products of proportions.
- A discussion of the potential bias due to "late-starters" in the development of cumulative detection probability as a function of range. Recommendations for eliminating (or at least minimizing) this bias are included in this report.

## TABLE OF CONTENTS

	<u>Page</u>
I. INTRODUCTION .....	1
II. SYMMETRIC CONFIDENCE INTERVALS FOR SEARCH RATE AND MEAN TIME-TO-DETECTION FOR THE ACTIVE AND PASSIVE AREA SEARCH SCENARIOS .....	4
III. APPROXIMATE CONFIDENCE INTERVALS FOR CUMULATIVE DETECTION PROBABILITY CURVES .....	18
IV. CONFIDENCE INTERVALS FOR PRODUCTS OF PROPORTIONS .....	29
V. A NOTE ON THE EFFECT OF "LATE STARTERS" ON THE ESTIMATE OF CUMULATIVE DETECTION PROBABILITY .....	44
REFERENCES .....	51
DISTRIBUTION LIST .....	52

## I. INTRODUCTION

Analysis & Technology, Inc. has been conducting a research study on statistical methods for the design and testing of submarine exercise data. This research has been sponsored by NAVAL ANALYSIS PROGRAMS (Code 462) of the OFFICE OF NAVAL RESEARCH under Contract Number N00014-72-C-0238.

The purpose of this study is to provide a basis for improvement in the design, analysis, and evaluation of submarine exercises and the exercise results through the development and application of statistical techniques. The results of this study are to be included in a revised edition of the SUBMARINE ANALYSIS NOTEBOOK (reference (1)). The notebook will contain instructions, procedures, standardized tests, and analytical techniques to evaluate in advance the data plans and requirements for proposed submarine exercises, and a description of the post-exercise statistical testing and general analysis necessary for the evaluation of the recorded data.

Although the SUBMARINE ANALYSIS NOTEBOOK will be the final product of this research and will utilize all of the results of this study, it is intended to be a convenient user's guide and not a compendium of statistics. Its purpose is to provide sufficient theory to allow the user to employ the statistical techniques correctly and to understand the implications of the tests. Therefore, certain theory and rationale behind the various statistical

tests will not be included in the Notebook. Instead, the user is referred to the appropriate textbook(s) and published papers for a detailed explanation of the theory.

Certain results of the research conducted under this contract merit separate publication since they require new or non-standard techniques. These results are presented in this methodology paper. The treatment of the techniques discussed herein is analytical rather than computational.

Chapter II of this report presents a method for calculating confidence intervals for search rate and mean time-to-detection for the active and passive area search scenarios. This method has application in both the design phase of an exercise and in the post-exercise analysis phase. In the planning stages, the test designer can use it to estimate sample size requirements for a desired confidence interval around the sample measure, or conversely, he can estimate the confidence interval he can expect to obtain from a predetermined sample size. In the post-exercise analysis phase, the analyst can apply the technique to calculate exact confidence intervals around the sample measure of search rate and mean time-to-detection.

In Chapter III, a methodology is given for obtaining approximate confidence intervals on cumulative detection probability (CDP) as a function of range. Considerable research was devoted to this problem before a

simple, practical method could be developed. The complexity of the problem is due to the inclusion of "turn-arounds" or CPAs and "late starters" in the construction of the estimate of the CDP curve.

Chapter IV discusses various methods of establishing approximate confidence intervals on products of proportions with respect to their usefulness in ASW applications.

The method due to Madansky (reference (2)) is recommended for inclusion in the SUBMARINE ANALYSIS NOTEBOOK since it appears to be most applicable to the type of exercise data used for obtaining estimates of Measures of Effectiveness (MOEs) for submarine missions.

The research conducted on cumulative detection probability (CDP) as a function of range revealed that the standard estimate of this function is susceptible to bias under certain conditions. The results of an investigation of this potential bias are presented in Chapter V.

II. SYMMETRIC CONFIDENCE INTERVALS FOR SEARCH RATE AND MEAN TIME-TO-DETECTION FOR THE ACTIVE AND PASSIVE AREA SEARCH SCENARIOS

Discussion

A primary measure of interest for the area search mission is search rate as calculated from exercise data. Search rate is defined as the rate at which the ASW unit searches its area, expressed in area per unit time (for example, square nautical miles per hour). A related measure is mean time-to-detection, which is inversely proportional to search rate. These measures are applicable to both active and passive area search scenarios and are discussed in detail in the current SUBMARINE ANALYSIS NOTEBOOK (reference (1)).

In planning an area search exercise, the test designer needs to investigate sample size requirements for obtaining statistically valid estimates of search rate and mean time-to-detection. This is true whether or not he has control over the sample size. If he does have control, then he can decide how many runs to schedule in order to obtain a desired confidence interval around the exercise estimates of these measures. If he does not have control over the sample size (i.e., the sample size has been fixed prior to the design of the exercise), then he can estimate, before the exercise, the anticipated confidence interval around the sample estimates of search rate and mean time-to-detection. In either case, the designer has a tool for assessing the statistical validity of the exercise results, prior to the conduct of the exercise.

After the exercise, when the data are used to calculate estimates of search rate and mean time-to-detection, the exercise analyst can calculate confidence intervals around his estimates, as a guide to their probable accuracy.

A method for determining exact symmetric confidence intervals on these measures has been derived from Koopman's (reference (3)) formulation of the area search problem, and is presented in this chapter.

The technique provides a functional relationship between confidence level, the width of the confidence interval as a percent of the sample estimate of the measures (search rate or mean time-to-detection), and the number of detections. It is the tool the analyst needs to make the pre-exercise decisions and to do the post-exercise analysis.

Since sample size is expressed in terms of detections and not in terms of exercise runs, the technique may require some prediction in the pre-exercise phase. Unless the test designer is able to specify that a test continue until a required number of detections has occurred, he needs a predicted value of search rate or mean time-to-detection in order to convert the required number of detections to the required number of runs.

### Development of Confidence Intervals for Search Rate

The cumulative detection probability as a function of time for a submarine conducting an area search is presented in the SUBMARINE ANALYSIS NOTEBOOK (reference (1)) as:

$$CDP = 1 - \exp (-\lambda t) \quad (\text{II.1})$$

where:

$$\lambda = \frac{1}{\bar{t}} = \frac{D}{\sum_{i=1}^N t_i},$$

$\bar{t}$  = sample mean time-to-detection ( $\bar{t} = \frac{1}{\lambda}$ ),

D = number of targets detected by the searcher,

$t_i$  = the length of the  $i$ th time interval of target exposure, ordered so that  $t_i \geq t_{i-1}$ ,

N = total number of time intervals,  $t_i$ .

Further, the quantity search rate (SR) is defined as:

$$SR = \frac{DA}{\sum_{i=1}^N t_i} = A\lambda = \frac{A}{\bar{t}} \text{ (area per unit time)} \quad (\text{II.2})$$

where:

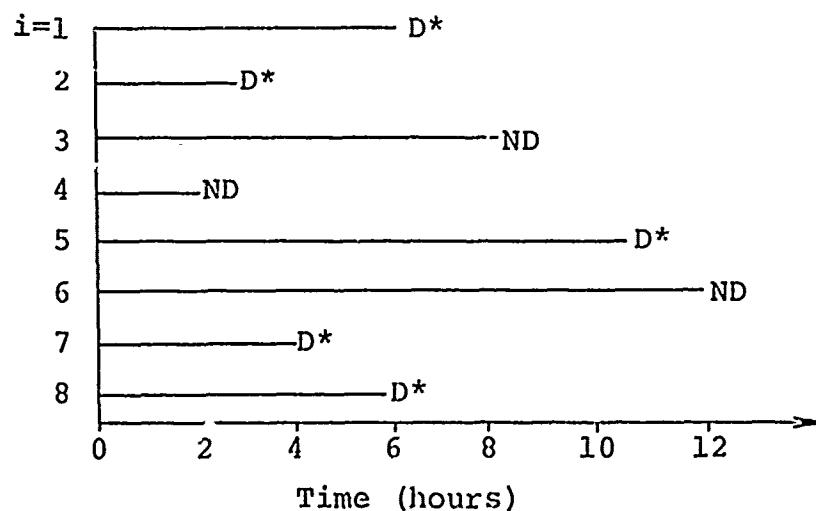
A = size of the search area in square nautical miles.

The quantity  $\bar{t}$  needs to be rewritten in a form which is consistent with the usual definition of a mean; i.e.,

$$\bar{x} = \frac{1}{M} \sum_{i=1}^M (x_i).$$

Referring to Dr. B. O. Koopman's Search and Screening (reference (3)), the probability  $\lambda dt$  of detecting in a short time interval of length  $dt$  is independent of time; in fact, the probability of a detection occurring in any short time interval of length  $dt$  is constant. Thus, it is possible to add a length of time in which no detections occurred to any other time interval, and especially to a time interval which ended in a detection.

For example, let us consider an area search exercise which produced the following data. The D\* stand for detections and the ND stand for no detections,

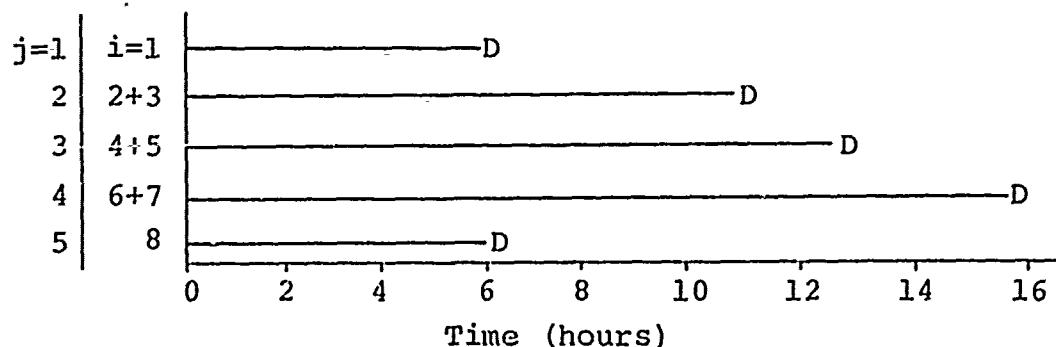


where:  $t_1 = 6$  hours,  $t_4 = 2$  hours,  $t_7 = 4$  hours,  
 $t_2 = 2.5$  hours,  $t_5 = 10.5$  hours,  $t_8 = 6$  hours,  
 $t_3 = 8$  hours,  $t_6 = 12$  hours,  $D = 5$ .

Runs number 1, 2, 5, 7, and 8 ended in detection, while the remaining three runs ended before the target was detected. In this case, the mean time-to-detection ( $\bar{t}$ ) can be calculated using equation (II.1),

$$\bar{t} = \frac{1}{D} \sum_{i=1}^N t_i = \frac{1}{5} \sum_{i=1}^8 t_i = 10.2.$$

However, since the probability of detection in a small time interval is constant, the data may be regrouped so that every time interval ends in a detection. A possible rearrangement of the data in this example is presented below:



where:  $t_1^* = 6$  hours,  
 $t_2^* = 10.5$  hours,  
 $t_3^* = 12.5$  hours,  
 $t_4^* = 16$  hours,  
 $t_5^* = 6$  hours.

Assume that  $t_1^*, t_2^*, \dots, t_D^*$  form a random sample from a population whose probability density function is

$$f(t) = \lambda^* e^{-\lambda^* t}, \quad 0 < t < \infty \quad (\text{II.3})$$

with the cumulative distribution function

$$F(t) = 1 - e^{-\lambda^* t}, \quad 0 < t < \infty. \quad (\text{II.4})$$

This follows from the premise that  $\lambda^*$  is assumed to be unknown,  $\lambda^* > 0$ , and is to be estimated by  $\lambda = 1/\bar{t}$ .

Define the random variable

$$X = 2\lambda^* t^*. \quad (\text{II.5})$$

Clearly,  $t^* = X/2\lambda^*$ ,  $dt^*/dX = 1/2\lambda^*$  and by the usual change of variable technique (reference (4)), the probability density function of  $X$  is

$$g(X) = (1/2)e^{(-1/2)X}, \quad 0 < X < \infty, \quad (\text{II.6})$$

the chi-square density function with 2 degrees of freedom. The values  $X_i = 2\lambda^* t_i^*$  ( $i = 1, \dots, D$ ) form a random sample from this chi-square density function. Thus, by the reproductive property of the chi-square,  $\sum_{i=1}^D X_i = \sum X_i = 2\lambda^* \sum t_i^*$  is distributed as a chi-square random variable with  $2D$  degrees of freedom. Let  $\chi_{\alpha}^2 (2D)$  and  $\chi_{1-\beta}^2 (2D)$  be respectively  $[100 \alpha]$  and  $[100 (1-\beta)]$  percentage points of the chi-square with  $2D$  degrees of freedom. We have

$$P(\chi_{\alpha}^2(2D) \leq \sum X_i \leq \chi_{1-\beta}^2(2D)) = 1 - (\alpha + \beta) = \gamma. \quad (\text{II.7})$$

But  $\sum X_i = 2\lambda^* \sum t_i^*$ . Hence

$$P(\chi_{\alpha}^2(2D) \leq 2\lambda^* \sum t_i^* \leq \chi_{1-\beta}^2(2D)) = 1 - (\alpha + \beta) = \gamma \quad (\text{II.8})$$

or equivalently

$$P(\chi_{\alpha}^2(2D)/2\sum t_i^* \leq \lambda^* \leq \chi_{1-\beta}^2(2D)/2\sum t_i^*) = 1 - (\alpha + \beta) = \gamma. \quad (\text{II.9})$$

The endpoints of the interval in equation (II.9) give a confidence interval on  $\lambda^*$  with confidence coefficient  $\gamma$ . Continuing, we can use the well known results

$$SR^* = A\lambda^* \text{ and } SR = \frac{DA}{\sum t_i} = \frac{DA}{\sum t_i^*} \quad (\text{II.10})$$

where  $SR^*$  = true (but unknown) search rate and  $SR$  denotes the sample estimate of the search rate. Substituting into (II.9) we have

$$P\left(\frac{SR\chi_{\alpha}^2(2D)}{2D} \leq SR^* \leq \frac{SR\chi_{1-\beta}^2(2D)}{2D}\right) = 1 - (\alpha + \beta) = \gamma, \quad (\text{II.11})$$

yielding a  $\gamma = [(1 - (\alpha + \beta)) 100\%]$  confidence interval on  $SR^*$ . The formulas

$$L_{\alpha} = \frac{SR\chi_{\alpha}^2(2D)}{2D} \text{ and } U_{1-\beta} = \frac{SR\chi_{1-\beta}^2(2D)}{2D} \quad (\text{II.12})$$

yield asymmetrical confidence intervals (about  $SR^*$ ) if  $\alpha = \beta$ . However, there is nothing inherently desirable about asymmetrical confidence limits. In fact, in pre-exercise design (i.e., determination of sample sizes or number of detections) symmetrical intervals are more desirable. Furthermore, in some cases the width of the asymmetric intervals are wider than the corresponding symmetrical ones. To obtain the symmetrical confidence intervals, find  $\alpha$  and  $\beta$  such that

$$\frac{SR\chi_{\alpha}^2(2D)}{2D} = (1-\Delta) SR = SR - \Delta(SR) \quad (\text{II.13})$$

and

$$\frac{SR\chi_{1-\beta}^2(2D)}{2D} = (1 + \Delta) SR = SR + \Delta(SR), \quad (\text{II.14})$$

where  $\Delta$  is the percent accuracy,  $0 < \Delta < 1$ .

That is

$$\chi_{\alpha}^2(2D) = 2D(1-\Delta) \quad (\text{II.15})$$

and

$$\chi_{1-\beta}^2(2D) = 2D(1+\Delta). \quad (\text{II.16})$$

Figure II.1, appearing at the end of this chapter, presents curves giving the relationship between the number of detections ( $D$ ) and the percent accuracy of the sample search rate ( $\Delta$ ), for four confidence levels ( $\gamma$ ); namely  $\gamma = 80\%$ ,  $90\%$ ,  $95\%$  and  $99\%$ . The application of these curves is illustrated in the examples presented in the last section of this chapter.

#### Development of Confidence Intervals for Mean Time-to-Detection

The above techniques can be adjusted to yield confidence intervals on  $1/\lambda^*$ , the "true" but unknown mean time-to-detection. From (II.9) we have

$$P\left(\frac{2D\bar{t}}{\chi_{1-\beta}^2(2D)} \leq \frac{1}{\lambda^*} \leq \frac{2D\bar{t}}{\chi_{\alpha}^2(2D)}\right) = 1 - (\alpha + \beta) \quad (\text{II.17})$$

where  $\bar{t} = \sum_{i=1}^N t_i/D = \sum_{i=1}^D t_i^*/D$ . Thus

$$L_{1-\beta} = 2D\bar{t}/\chi_{1-\beta}^2(2D) \quad (\text{II.18})$$

and

$$U_{\alpha} = 2D\bar{t}/\chi_{\alpha}^2(2D) \quad (\text{II.19})$$

yield asymmetrical confidence intervals on  $1/\lambda^*$ . To obtain symmetrical confidence intervals on  $1/\lambda^*$ , set  $L = (1-\Delta)\bar{t}$  and  $U = (1+\Delta)\bar{t}$ . Solving these equations, we have

$$\chi^2_{1-\beta}(2D) = 2D/(1-\Delta), \quad (II.20)$$

$$\chi^2_{\alpha}(2D) = 2D/(1+\Delta). \quad (II.21)$$

Again, by fixing  $D$ , one can adjust  $\Delta$ ,  $\alpha$  and  $\beta$ , subject to  $\gamma = [1-(\alpha+\beta)] = 99\%$  (or 95%, 90%, 80%), until equations (II.20) and (II.21) are satisfied.

Figure II.2 contains graphs giving the relationship between the number of detections ( $D$ ), and the percent accuracy of the sample mean time-to-detection ( $\Delta$ ), for four values of  $\gamma = [1-(\alpha+\beta)]$ , namely  $\gamma = 80\%$ ,  $90\%$ ,  $95\%$  and  $99\%$ .

#### Application

The following examples illustrate the use of Figures II.1 and II.2 in pre and post-exercise analysis.

#### Example 1 - (Pre-Exercise)

An active area search exercise is being designed to measure active search rate. The test designer wishes to estimate the number of detections needed to be 90% confident that a  $\pm 20\%$  interval around the sample search rate contains the true, but unknown, value of search rate.

In this case, the designer will not attempt to predict the value of the sample search rate since the effects of target forestalling (i.e., target

counterdetecting first and then avoiding detection) may be significant. However, his estimate of the required sample size must be in terms of detections and not in terms of exercise runs.

For this problem, the following values of confidence level and percent accuracy are used:

Confidence Level = 90%

Percent Accuracy = 20%

Confidence Limit =  $SR \pm .2SR$

Using Figure II.1, the required number of detections is 65.

#### Example 2 - (Pre-Exercise)

In designing a passive area search exercise, the test designer learns that the proposed exercise will consist of 60 runs, each 24 hours long. He wishes to estimate the confidence intervals around the sample mean time-to-detection, prior to the conduct of the exercise.

From prediction and/or prior passive area search exercises, the test designer estimates that the cumulative detection probability at the end of each run is .87. Thus, the estimated number of detections at the end of 60 runs is approximately 52 (i.e.,  $.87 \times 60$ ).

The predicted, pre-exercise value of mean time-to-detection is approximately 28 hours ( $\bar{t} = \text{total search time}/D \approx 27.69$ ).

Using Figure II.2, he obtains the following estimates of three confidence intervals around the sample mean time-to-detection based on  $D = 52$  and  $\bar{t} = 27.69$ :

<u>Confidence Level</u>	<u>Percent Accuracy, <math>\Delta</math></u>	<u>Confidence Interval (<math>\bar{t}-\Delta\bar{t}</math>, <math>\bar{t}+\Delta\bar{t}</math>)</u>
80%	.18	(22.71, 32.67)
90%	.23	(21.40, 33.98)
95%	.28	(19.94, 35.44)

Example 3 - (Post-Exercise)

The analyst wishes to compute 80%, 90%, 95% and 99% confidence intervals around the exercise (sample) value of search rate. The sample search rate was calculated to be 23 square nautical miles per day based on 45 detections.

Using Figure II.1, he obtains the following values of the exact symmetric confidence intervals around the exercise search rate.

<u>Confidence Level</u>	<u>Percent Accuracy, <math>\Delta</math></u>	<u>Confidence Interval (<math>SR-\Delta SR</math>, <math>SR+\Delta SR</math>)</u>
80%	.19	(18.63, 27.37)
90%	.24	(17.48, 28.52)
95%	.29	(16.33, 29.67)
99%	.38	(14.26, 31.74)

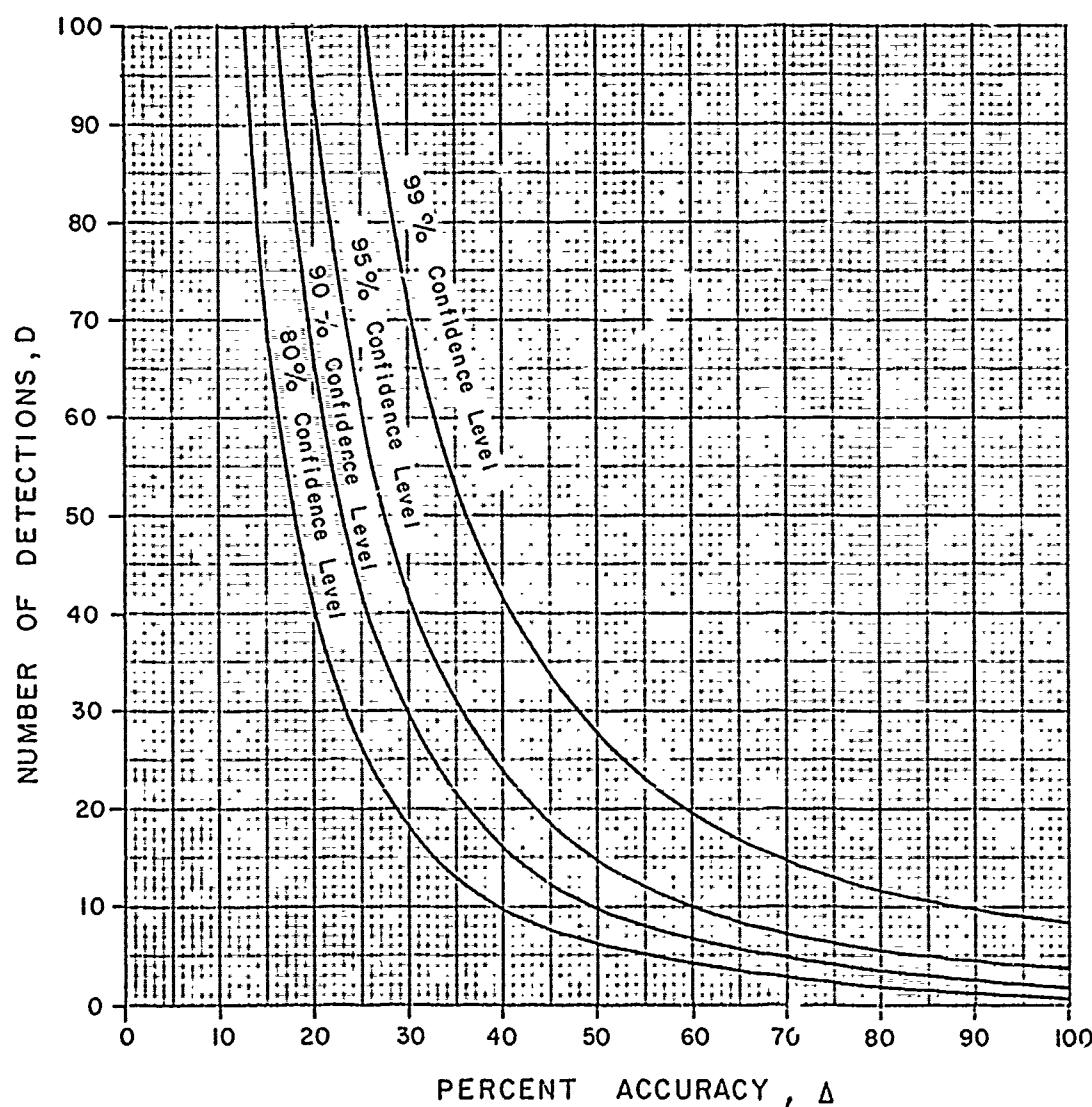


Figure II.1 Relationship between number of detections,  $D$ , and percent accuracy,  $\Delta$ , for 80%, 90%, 95% and 99% confidence intervals on true search rate,  $SR^*$ . For  $D$  detections, the limits of the interval are  $SR \pm \Delta SR$ , where  $SR =$  sample search rate.

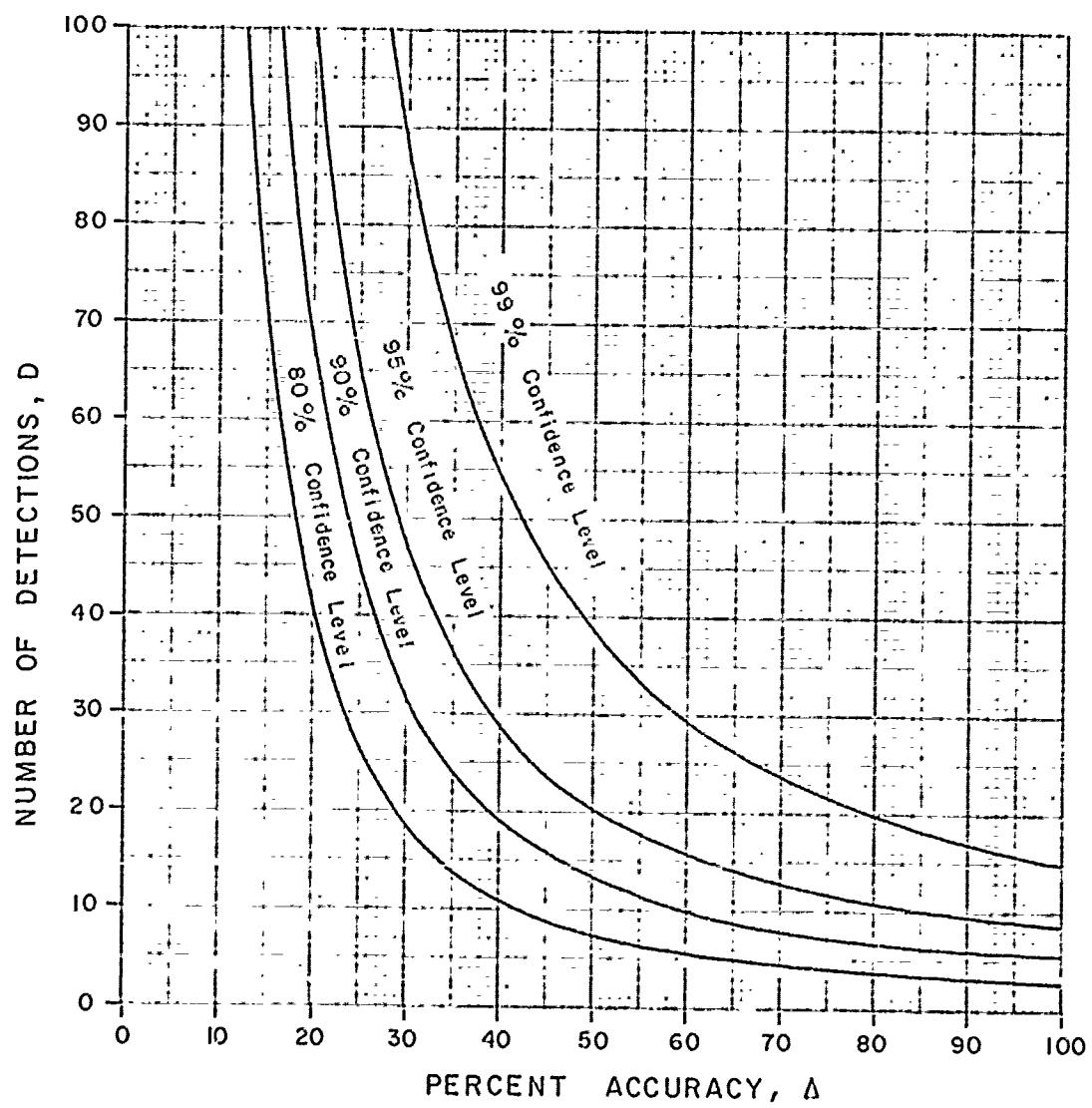


Figure II.2 Relationship between number of detections,  $D$ , and percent accuracy,  $\Delta$ , for 80%, 90%, 95% and 99% confidence intervals on true mean time-to-detection,  $1/\lambda^*$ . For  $D$  detections the limits of the interval are  $\bar{t} \pm \Delta\bar{t}$ , where  $\bar{t}$  = sample mean time-to-detection.

### III. APPROXIMATE CONFIDENCE INTERVALS FOR CUMULATIVE DETECTION PROBABILITY CURVES

#### Discussion

Cumulative detection probability as a function of range ( $CDP = f(R)$ ) is an important performance measure of a sonar system. As developed from exercise data, it provides the analyst with an estimate of the system's detection performance in terms of the probability of detecting a target by the time the target has closed to within a specified range.

The development of  $CDP = f(R)$  is simple if each target closed until it was detected or until it reached (approximately) zero range from the detecting unit. In this case, when  $N$  trials have been made, the cumulative detection probability as a function of range can be determined as:

$$CDP(R) = \frac{D_R}{N}$$

where:

$D_R$  = Number of detections made at range  $R$   
or greater.

However, in actual exercises, the target does not continue to close indefinitely until it is detected. Even if it were required to do so, the resultant CDP

curve would not be representative of real targets that are free to maneuver. Thus, a realistic target may reach a closest point of approach (CPA) after which it begins opening range. In many cases, such a target may never be detected during the run; and further, the observed CPA of a run may be at a lesser range than several of the detection ranges of other runs. These undetected targets are called CPAs or turn-arounds.

In addition, a typical exercise may include targets that were "late-starters". These late-starters became detection opportunities at a lesser range than some or many of the detection ranges of other targets.

A detailed discussion of the development of cumulative detection probability as a function of range for the general case involving CPAs and late-starters is presented in the current SUBMARINE ANALYSIS NOTEBOOK (reference (1)). While several forms of the CDP equation are presented in the Notebook, the following equation is the most useful for our development of confidence intervals:

$$CDP_i = 1 - \prod_{j=1}^i g_j$$

where:

$CDP_i$  = The cumulative detection probability at range  $i$ .

$g_j$  = The probability of no-detection in the  $j^{\text{th}}$  range band, on a target which was not detected before entering the  $j^{\text{th}}$  range band.

## Methodology

Several techniques (e.g., references (5) and (6)) have been proposed in attempts to solve, at least approximately, the problem of obtaining confidence intervals on cumulative detection probability (CDP) as a function of range. In the simple case with no CPAs or late-starters, confidence limits can be calculated using standard techniques since the function

$$CDP(R) = \frac{D_R}{N}$$

is, at each range, an observation on a binomial population. Specifically, there is a probability (equal to  $CDP(R)$ ) that a closing target will be detected at a range greater than or equal to  $R$ . The observed fraction represents  $D_R$  successes in  $N$  trials.

In contrast, consider the more complicated case where CPAs and late-starters are included in the data. Suppose a detection occurs at some range  $R^*$  that is less than the longest detection range, and less than at least one CPA range or late-starter starting range. It is not possible to characterize  $CDP(R^*)$  as  $D$  successes in  $N$  trials since not all the runs that were valid trials at the starting range are still valid trials at  $R^*$  (due to CPA's), and not all the runs that are valid trials at  $R^*$  have been valid trials over the entire interval from the starting range to  $R^*$  (due to late-starters). In short, although it is still possible to count the successes, it is no longer possible to count the number of trials; and hence, the binomial confidence interval technique is not directly applicable.

The method described here is designed to choose a suitable number to use as this unknown number of trials so that approximate confidence limits can be calculated using the binomial distribution. The choice is made in such a way that the resulting confidence intervals reflect, in a reasonable way, the actual sample size.

Consider the following notation:

1.  $R_i$  = range of an  $i^{\text{th}}$  detection, ordered so that  $R_{i+1} \leq R_i \leq R_{i-1}$ .
2.  $M_i$  = number of targets available at a range just less than range  $R_i$ .

As referred to in the Discussion, the equation for CDP given in the SUBMARINE ANALYSIS NOTEBOOK (reference (1)) is:

$$CDP_i = 1 - \prod_{j=1}^i g_j$$

Since the estimator for each  $g_j$  can be written in the form  $g_j = \frac{M_j}{M_j+1}$ , the  $CDP_i$  can be rewritten in the notation of this report.

3.  $CDP_i = 1 - \prod_{j=1}^i \frac{M_j}{M_j+1}$  = cumulative detection probability at range  $R_i$ .

4.  $N_i = \frac{i}{CDP_i}$  = estimated sample size necessary for producing  $i$  detections when the cumulative detection probability is  $CDP_i$ .
5.  $R_i^* =$  an arbitrary range satisfying  $R_{i+1} < R_i^* < R_i$ .
6.  $A_i^* =$  number of "late-starters" with starting range less than  $R_i$  but greater than or equal to  $R_i^*$ .
7.  $C_i^* =$  number of "turn arounds" where the closest point of approach (CPA) is less than  $R_i$  but greater than or equal to  $R_i^*$ .
8.  $k_i^* = C_i^* - A_i^* =$  net loss in number of targets (opportunities) between range  $R_i$  and  $R_i^*$ .
9.  $[N_i] =$  largest integer less than or equal to  $N_i = i/CDP_i$ .

Given a sample size  $N_i$  and a value of  $CDP_i$  for a range  $R_i$ , we could calculate the expected number of detections at or before  $R_i$  by using the following equation:

$$i = CDP_i \cdot N_i$$

In this case, since  $i$  and  $CDP_i$  are known, we can use the same equation to estimate the "effective sample size"

$$N_i = \frac{i}{CDP_i} .$$

There are two cases which arise in placing confidence intervals on the value of  $CDP_i$  at range  $R_i$ . In case 1,  $N_i$  is an integer. Here  $CDP_i = i/N_i$  can be treated as  $i$  detections, at ranges greater than or equal to  $R_i$ , out of  $N_i$  (estimated) opportunities. Hence, using the tables and charts for "exact" confidence intervals on a single proportion (see, for example, reference (1)) one can obtain an approximate  $\gamma\%$  confidence interval on  $CDP_i$ , where  $L_i$  and  $U_i$  denote, respectively, the lower and upper limits of the interval.

In case 2,  $N_i$  is not an integer. Here

$$i/([N_i] + 1) < CDP_i < i/[N_i]. \quad (\text{III.1})$$

Using the tables and charts for confidence intervals on a single proportion, we can obtain an approximate, but probably conservative,  $\gamma\%$  confidence interval on  $CDP_i$ . For the lower limit,  $L_i$ , take the lower limit of a  $\gamma\%$  confidence interval on the proportion  $i/([N_i] + 1)$ , i.e., treat  $CDP_i$  as  $i$  detections out of  $[N_i] + 1$  opportunities. For the upper limit,  $U_i$ , take the upper limit of a  $\gamma\%$  confidence interval on the proportion  $i/[N_i]$ , i.e., treat  $CDP_i$  as  $i$  detections out of  $[N_i]$  opportunities.

The above technique provides confidence intervals on CDP at those ranges where detections occurred. Consider range  $R_i^*$  where  $R_{i+1} < R_i^* < R_i$ , and recall that the estimated CDP for range  $R_i^*$  is still  $CDP_i$ . First, consider the case where  $A_i^* \geq C_i^*$  (i.e., the number of late-starters is greater than or equal to the number of turn arounds in

the interval from  $R_i$  to  $R_i^*$ ). For this case, no change is recommended in the confidence interval on CDP. As  $R_i^*$  moves from  $R_i$  to  $R_i + 1$ , there will be no change in the limits of the confidence interval so long as  $A_i^* \geq C_i^*$ .

Secondly, consider the case where  $A_i^* < C_i^*$  (the turn-arounds since the last detection outnumber the late-starters since the last detection). Intuitively, there is a drop in the precision of our estimate of CDP at range  $R_i^*$ . This drop in the precision should be expressed by calculating a wider confidence interval on  $CDP_i$  at range  $R_i^*$ . Further, it should be done in such a way that the upper limit of the interval goes to 100% if the number of opportunities,  $M_i - (C_i^* - A_i^*)$ , goes to zero. Since CDP is a monotonically decreasing function of range, we can still use  $L_i$  as the lower limit on  $CDP_i$  at range  $R_i^*$ .

As an intermediate step in the adjustment of the upper limit, compute an estimate of CDP, say  $CDP_i^*$ , as if the  $i^{th}$  detection occurred at range  $R_i^*$ .

We obtain

$$CDP_i^* = 1 - \left( \prod_{j=1}^{i-1} \frac{M_j}{M_j + 1} \right) \cdot \left( \frac{M_i - k_i^*}{M_i + 1 - k_i^*} \right) \quad (\text{III.2})$$

where  $k_i^* = C_i^* - A_i^*$ . The estimated number of opportunities producing this value of CDP in  $i$  detections is  $N_i^* = i/CDP_i^*$ . Of course our estimate of CDP at range  $R_i^*$  continues to be

$CDP_i$ , but for a (conservative) upper limit we take the upper limit of a  $\gamma\%$  confidence interval as if we have had  $i$  detections out of  $[N_i^*]$  opportunities. Note that if  $k_i^* = M_i$ , then there are no targets available at range  $R_i^*$  and  $CDP_i^* = 1$ . Hence, the upper limit on  $CDP_i$  at range  $R_i^*$  will be 1. This technique also guarantees that the upper limits of the confidence intervals will converge to 1 as the range goes to 0.

Using the above technique, it may happen that the upper limits of the intervals are not a monotonically decreasing function of range. See Figure III.1 for a sketch of an artificial example.

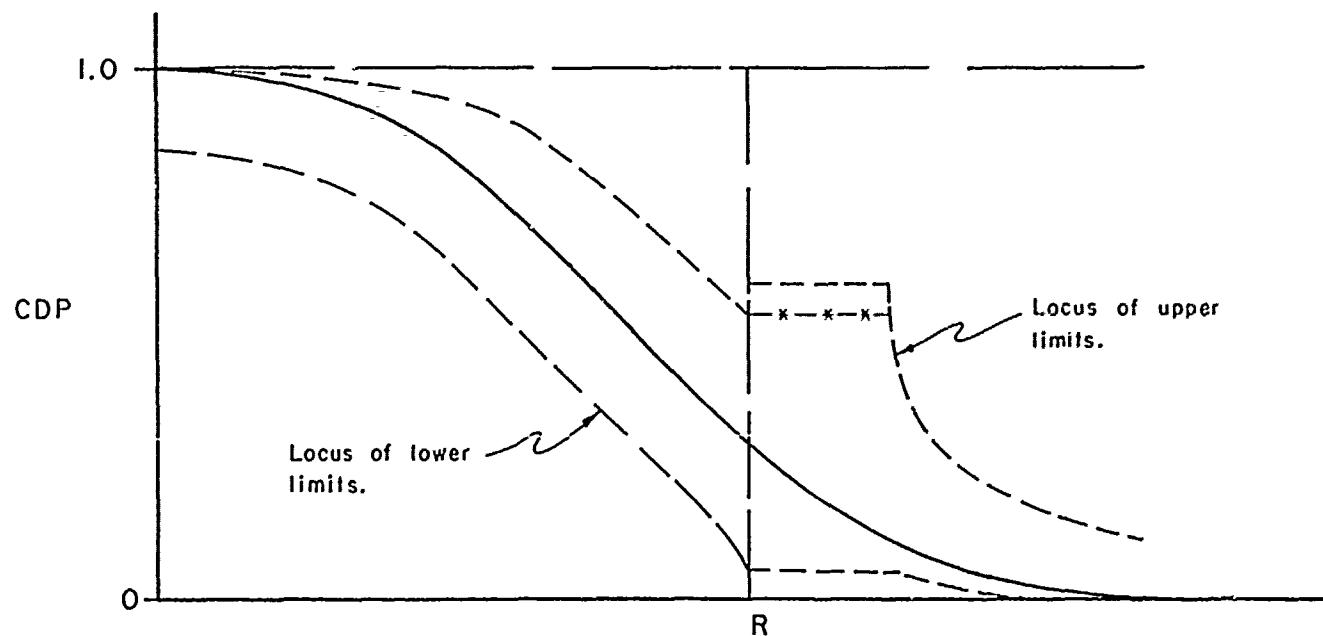


Figure III.1 A Situation Where the Locus of the Upper Limits of the Confidence Intervals on CDP is not a Monotonic Function of Range

Non-monotonic changes in the upper limits are possible if large numbers of late-starters are introduced to a relatively small number of opportunities. One may make the locus of the upper limits monotonically decreasing as the range increases. This adjustment is indicated by the symbols "-\*-\*-\*" in Figure III.1. This appears to be a reasonable adjustment of the upper limits since the "true" CDP is a monotonically decreasing function of range.

The data for the numerical example (fictitious) appearing in Table III.1 is taken from the current SUBMARINE ANALYSIS NOTEBOOK (reference (1)). Approximate 80% confidence limits on the CDP curve have been calculated and are included in the table. Note that because of the small number of opportunities and detections, the intervals are quite wide. The results are presented graphically in Figure III.2.

Detection Number	Detection Range	CPA, This Target for No-Detection	Start Range This Target	CDP <sub>i</sub>	N <sub>i</sub>	CDP <sub>i</sub> *	N <sub>i</sub> *	Approximate 80% Confidence Limits	
								L <sub>i</sub>	U <sub>i</sub>
0				0	5.			0.0	.275
1	12		20	1-.8=.2	5.			.021	.584
2	10		11	24		1-3/4=.25	4.	.021	.679
				14	1-.5333	4.3 =.4667			
			9	32		1- (.8) (1/2) = .6	3.3	.112	.966
3	7			8	1-.4267	5.2 =.5733		.201	.888
			6	8		1- (.5333) (3/4) = .6	5.0	.201	.888
4	5			8	1-.2845	5.6 =.7155		.333	.979
5	4			8	1-.1423	5.8 =.8577		.489	1.000
			3	27			No adjustment: Late starter at 3.5		
			2	3.5			$1 - (.1422) (0/2)$ = 1.0	5.0	.489 1.000

Table III.1 Computation of Cumulative Probability as a Function of Range with (Approximate) 80% Confidence Limits.

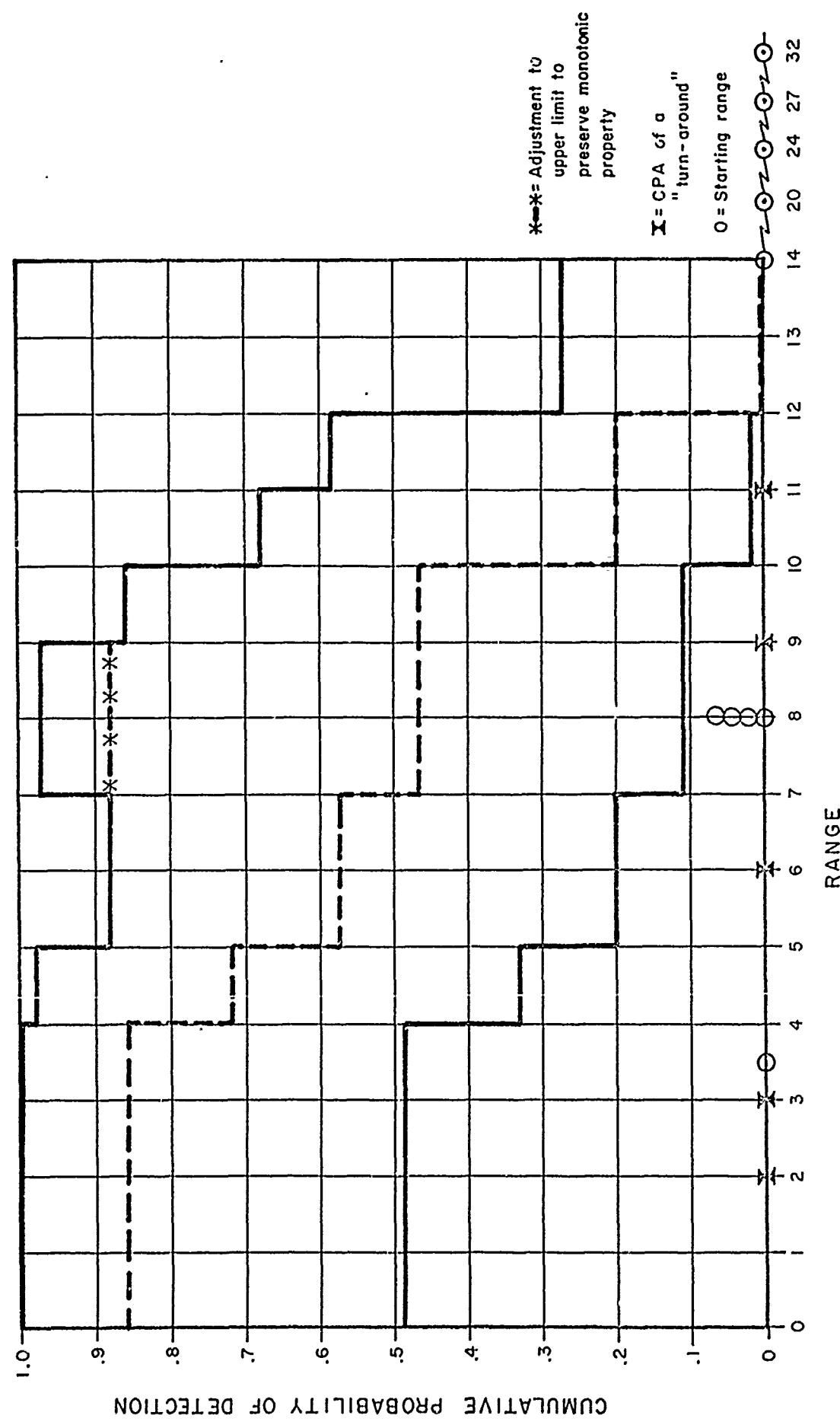


Figure III-2 A Sample Cumulative Probability of Detection as a Function of Range on a Closing Target With (Approximate) 80% Confidence Intervals

#### IV. CONFIDENCE INTERVALS FOR PRODUCTS OF PROPORTIONS

##### Discussion

It is well known that several Measures of Effectiveness (MOEs) useful in the analysis of Anti-Submarine Warfare exercises can be written in the form

$$MOE = p_1 p_2 \dots p_N = \prod_{i=1}^N p_i \quad (IV.1)$$

where  $p_i$  is the probability of success of the  $i$ th component of a system, given that the first  $(i-1)$  components have succeeded. In statistical terms, the MOE is a measure of the reliability of a series system (i.e., a system in which every component must succeed in order for the system to succeed) and is the probability that the system will succeed on a given "trial".

In this chapter, the problem of obtaining (approximate) confidence intervals on MOEs is considered. A brief description of techniques due to Harris (reference (7)), Madansky (reference (2)) and Walsh (reference (8)) is given, with a discussion of the conditions under which each is applicable. In addition, an in-depth review of a modified Bayesian technique is presented, along with some of the problems associated with its use. Tables are given comparing the various techniques under different conditions.

Based on the comparative analysis, the method due to Madansky is recommended for inclusion in the updated SUBMARINE ANALYSIS NOTEBOOK. It appears to be most applicable to the type of exercise data used for obtaining estimates of Measures of Effectiveness for submarine missions.

### Theory

Assume  $X_1, \dots, X_N$  are statistically independent random variables and for each  $i$ ,  $X_i$  has a binomial distribution with parameters  $n_i$  and  $p_i$ . As usual,  $X_i$  will be the number of "successes" out of  $n_i$  "trials" where the true, but unknown, probability of success on any trial is  $p_i$ .

Harris gives a brief review of prior work in this area and extends a general technique developed by Buehler (reference (9)) for obtaining approximate confidence intervals whenever each  $X_i$ ,  $i=1, \dots, N$ , is approximately Poisson distributed, (i.e., roughly  $n_i > 40$  with  $p_i < 5/n_i$ ). Thus,  $p_i$  must be "small" which restricts the use of the technique even in the case of "large" samples.

Walsh succeeded in deriving a function of  $\hat{p}_i = \bar{x}_i/n_i$ ,  $i = 1, \dots, N$  which can be "inverted" to obtain an approximate confidence interval on  $\prod_{i=1}^N p_i$ . The function is approximately normally distributed whenever the numbers of trials,  $n_i$ 's, are moderately large and the success probabilities,  $p_i$ 's, are of at least moderate size (i.e., roughly,  $p_i \geq 1/2$ ,  $n_i p_i \geq 10$ , and  $n_i (1-p_i) \geq 5$ ; or  $p_i \geq 2/3$ ,  $n_i p_i \geq 10$ , and  $n_i (1-p_i) \geq 2$ ). As before, the restrictions on  $p_i$  limit the use of the procedure.

Madansky derives approximate confidence intervals on  $\prod_{i=1}^N p_i$  by "inverting" the generalized likelihood ratio test and using the well known asymptotic Chi-square random variable. The technique gives good results whenever the  $n_i$ 's are "large" and the  $p_i$ 's are not close to zero (i.e., roughly,  $n_i > 30$  with  $p_i > 5/n_i$ ).

The above are "relatively large sample" techniques. Furthermore, each has restrictions on  $p_i$ ,  $i=1, \dots, N$ . In the case when these restrictions are not met there are no procedures, known to the authors, for obtaining confidence intervals on  $MOE = \prod_{i=1}^N p_i$ .

In an attempt to provide guidance in all cases, an investigation was made into the Bayesian confidence intervals suggested by Springer and Thompson (reference (10)). Under the assumption that  $p_i$ ,  $i=1, \dots, N$  has a uniform prior distribution on the interval from zero to one, they derived the Bayes posterior distribution of the MOE, conditional on the observed values  $\hat{p}_i = x_i/n_i$ . The appropriate percentage points of the posterior distribution form the limits of the Bayesian confidence interval on the MOE. Unfortunately, the Bayesian procedure has two undesirable features. First, the mean of the posterior distribution is

$$\hat{\mu} = \prod_{i=1}^N \frac{x_i + 1}{n_i + 2} \quad (IV.2)$$

Equation (IV.2) is a biased estimate of the MOE. As the authors point out, the estimate is unbiased if the  $n_i$  all tend to infinity. However, for moderate sample sizes,  $n_i$ ,

the bias can be serious if  $N$  is large and  $p_i$  is close to 1. See Table IV.1 for some simple numerical examples.

Table IV.1

Correct Estimate	Bayesian Estimate (Biased)
$\frac{N}{\prod_{i=1}^N (x_i/n_i)}$	$\frac{N}{\prod_{i=1}^N (x_i+1)/(n_i+2)}$
$\frac{5}{\prod_{i=1}^5 (\frac{19}{20})} = .7738$	$\frac{5}{\prod_{i=1}^5 (\frac{20}{22})} = .6209$
$\frac{10}{\prod_{i=1}^{10} (\frac{19}{20})} = .5987$	$\frac{10}{\prod_{i=1}^{10} (\frac{20}{22})} = .3855$
$\frac{15}{\prod_{i=1}^{15} (\frac{19}{20})} = .4633$	$\frac{15}{\prod_{i=1}^{15} (\frac{20}{22})} = .2394$

A second problem with the Bayesian limits is that for small sample sizes ( $n_i$ 's) they are too narrow in at least one case. Namely, when there is only one term in the product, i.e., the MOE is a single proportion. For this case, "exact" confidence intervals can be given as in

Table IV.2. Comparison of the Bayesian limits and the "exact" limits on a single proportion lead us to modify the Bayesian technique.

Table IV.2 80% Confidence Interval on a Single Proportion  
When the Sample Size = 5

Number of Successes X	Proportion of Successes	Bayesian Limits		Exact Confidence Limits	
		Lower	Upper	Lower	Upper
0	0.0	.019	.177	0.00	.275
1	.2	.098	.495	.021	.584
2	.4	.209	.656	.112	.753
3	.6	.343	.790	.247	.888
4	.8	.496	.902	.416	.979
5	1.0	.822	.981	.725	1.000

In the Bayesian spirit, assume that for each  $i$ ,  $p_i$  is a random variable. Given  $\hat{p}_i = x_i/n_i$ , the limits  $(L_i, U_i)$  of an exact  $1-\alpha$  confidence interval on  $p_i$  are known (see, for example, reference (11)) to satisfy

$$\text{BETA}(L_i; x_i, n_i - x_i + 1) = \alpha/2 \quad (\text{IV.3})$$

and

$$\text{BETA}(U_i; X_i+1, n_i - X_i) = 1 - \alpha/2 \quad (\text{IV.4})$$

respectively. The function

$$\begin{aligned} \text{BETA}(z; A, B) &= \frac{\Gamma(A+B)}{\Gamma(A)\Gamma(B)} \int_0^z t^{A-1} (1-t)^{B-1} dt \\ &= \int_0^z f(t; A, B) dt \end{aligned} \quad (\text{IV.5})$$

is the cumulative form of the beta distribution. Given  $\hat{p}_i = X_i/n_i$  with  $0 < X_i < n_i$ , define the "interval generating function" of  $p_i$  to be

$$F_i(p_i; X_i, n_i) = \begin{cases} \text{BETA}(p_i; X_i, n_i - X_i + 1) & \text{if } 0 \leq \text{BETA}(p_i; X_i, n_i - X_i + 1) \leq .5 \\ \text{BETA}(p_i; X_i + 1, n_i - X_i) & \text{if } 1 \geq \text{BETA}(p_i; X_i + 1, n_i - X_i) \geq .5 \\ .5 & \text{if } \underline{p}_i \leq p_i \leq \bar{p}_i \end{cases} \quad (\text{IV.6})$$

where  $p_i$  and  $\bar{p}_i$  satisfy  $\text{BETA}(p_i; X_i, n_i - X_i + 1) = .5$  and  $\text{BETA}(\bar{p}_i; X_i + 1, n_i - X_i) = .5$  respectively.

A sketch of the graph of  $F(p;3,5)$  appears in Figure IV.1. When  $\hat{p} = 3/5$ , confidence intervals for  $p$  may be read directly from the graph. For example, with  $F(p;3,5) = .05$  and  $.95$  we obtain  $p = .189$  and  $.924$  respectively. The interval  $(.189, .924)$  is a 90% confidence interval on  $p$ .

If  $X_i = 0$ , define

$$F_i(p_i; 0, n_i) = \text{BETA}(p_i; 1, n_i); \quad 0 \leq p_i \leq 1. \quad (\text{IV.7})$$

If  $X_i = n_i$ , define

$$F_i(p_i; n_i, n_i) = \text{BETA}(p_i; n_i, 1); \quad 0 \leq p_i \leq 1. \quad (\text{IV.8})$$

Sketches of the graphs of  $F(p;0,5)$  and  $F(p;5,5)$  appear in Figure IV.2. In the case  $X_i = 0$  the lower limit,  $L_i$ , of the  $(1-\alpha)$  confidence interval on  $p_i$  is 0 and the upper limit satisfies  $F_i(U_i; 0, n_i) = 1-\alpha$ . For example, from Figure IV.2 we see that a 90% confidence interval on  $p$  when  $X = 0$ ,  $n = 5$  is  $(0, .369)$ . Similarly, if  $X_i = n_i$ , then the upper limit,  $U_i$  is taken to be 1 and the lower limit,  $L_i$ , satisfies  $F_i(p_i; n_i, n_i) = \alpha$ . From Figure IV.2, a 90% confidence interval on  $p$  is  $(.631, 1)$  when  $X = 5$ ,  $n = 5$ .

By this time it is obvious that we are requiring the percentage points of the interval generating function,  $F_i(p_i; X_i, n_i)$ , to be the limits of an exact confidence interval on  $p$ .

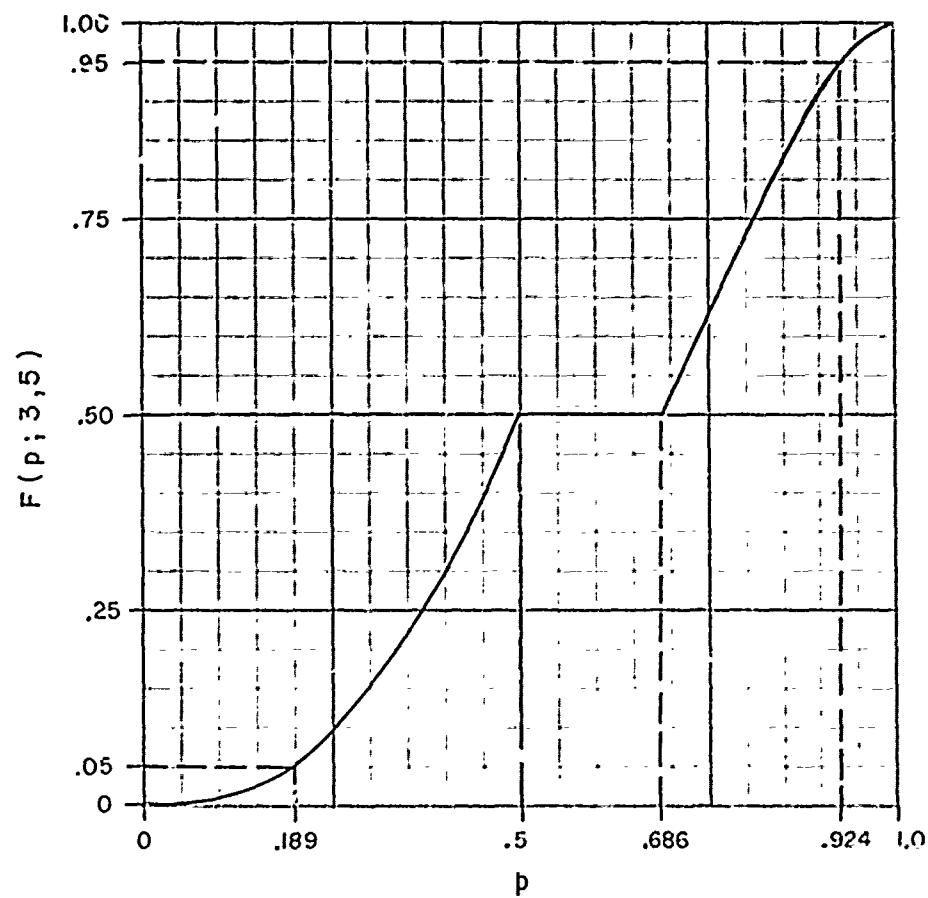


Figure IV.1 Graph of the interval generating function  $F(p; 3, 5)$  giving confidence intervals on  $p$  when  $\hat{p}=3/5$ .

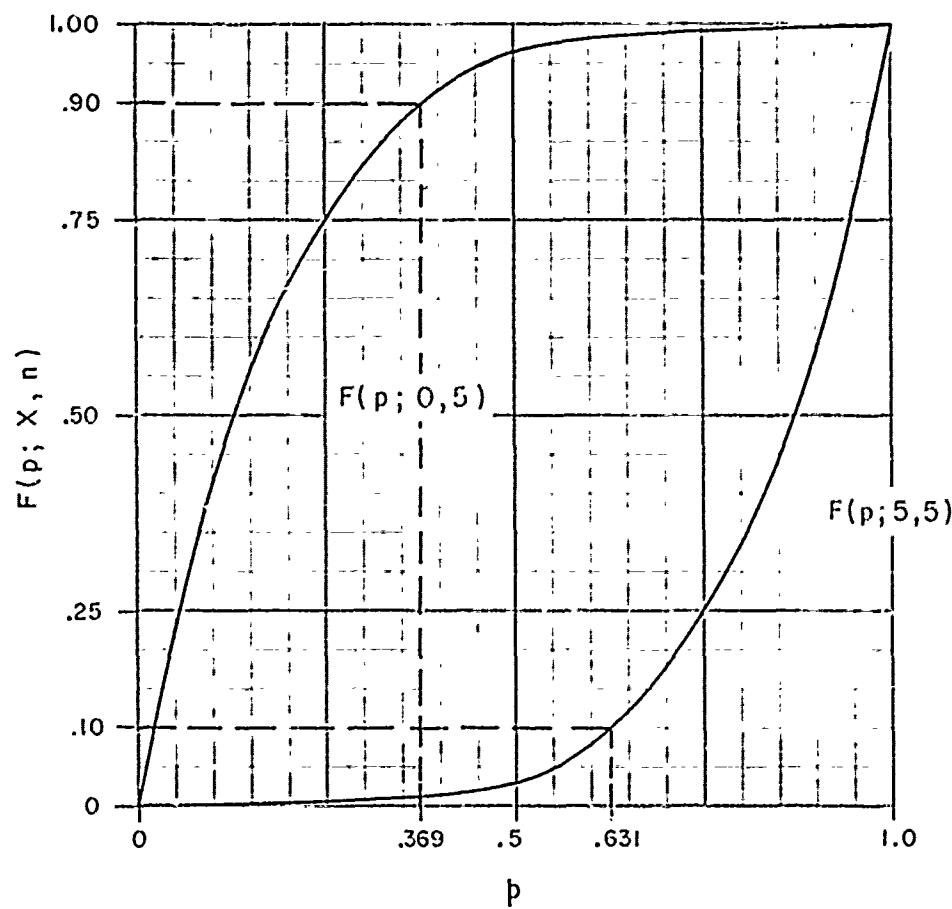


Figure IV.2 Graphs of the interval generating functions  $F(p; 0, 5)$  and  $F(p; 5, 5)$ .

Also, the function  $F_i(p_i; x_i, n_i)$  satisfies necessary and sufficient conditions (see, for example reference (12)) to be a cumulative probability distribution function. This motivates the following definition:

Definition 4.1. Conditional on the observed value  $\hat{p}_i = \bar{x}_i/n_i$ , the modified Bayesian "posterior" cumulative probability distribution of  $p_i$  is given by  $F_i(p_i; x_i, n_i)$  for  $i = 1, 2, \dots, N$ .

We desire the posterior cumulative probability distribution (cpd) of  $MOE = \prod_{i=1}^N p_i$  conditional on  $\prod_{i=1}^N \hat{p}_i$ . The posterior cpd is not available in closed form. However, it can be simulated easily on a high speed computer. The modified Bayesian confidence intervals will then be given by the appropriate percentage points in the simulated posterior cpd of  $MOE$ .

To simulate the posterior cpd of  $R$ , first generate uniformly distributed random numbers,  $r_i$ , between 0 and 1 and solve the equations

$$r_i = F_i(p_i; x_i, n_i); \quad i=1, \dots, N, \quad (IV.9)$$

for  $p_i$ . Call the solutions  $p_i^*$ ;  $i=1, \dots, N$ . This gives random observations from the posterior distributions of  $p_i$ ,  $(i=1, \dots, N)$ . Form the product  $R^* = \prod_{i=1}^N p_i^*$ .

Definition 4.2. The product  $R^* = \prod_{i=1}^N p_i^*$  is a random observation from the posterior cpd of the MOE.

Continuing in the above manner, generate  $M$  observations from the posterior cpd of the MOE, and denote them by

$R_1^*, R_2^*, \dots, R_M^*$ . Let  $PC_j$  denote the  $j$ th percentile of the  $R_i^*$ 's, i.e.,  $PC_j$  is the value such that  $j\%$  of the  $R_i^*$ 's are less than or equal to  $PC_j$ . Clearly, if  $M$  is large enough, the limits of the 80% modified Bayesian confidence interval on  $R$  will be given approximately by  $PC_{10}$  and  $PC_{90}$ . The limits of the 90% modified Bayesian confidence interval on  $R$  will be given approximately by  $PC_5$  and  $PC_{95}$ , etc.

A program was written in GE MK II Time-Sharing Fortran for obtaining the modified Bayesian confidence intervals and is available upon request.

The accuracy of the program yielding the modified Bayesian intervals depends on  $M$ , the number of simulations, and the accuracy of the internal computer function used to solve the incomplete beta function. The accuracy of the internal incomplete beta function available on our computer is not too good. In extreme cases, agreement with tabled values is to no more than 2 or 3 digits. This, coupled with the numerical technique used to fit the curve  $F_i(p_i; x_i, n_i)$  by a series of straight lines, did not yield the 2 or 3 digit accuracy hoped for. However, so long as  $M \geq 40/\alpha$ , the end points of the  $1-\alpha$  confidence interval on a single proportion ( $N=1$ ) were within  $\pm .01$  of the exact confidence intervals appearing in reference (11). See, for example, the figures in Table IV.3.

Table IV.3

Confidence Intervals on  $p$  when  $X=1, n=5$ 

Confidence ( $1-\alpha$ )	Exact Confidence Limits	Modified Bayesian Intervals			
		$M=100$	$M=200$	$M=400$	$M=1000$
80%	(.02,.58)	(.02,.57)	(.02,.56)	(.02,.57)	(.02,.58)
90%	(.01,.66)	(.010,.64)	(.01,.66)	(.01,.67)	(.01,.66)
95%	(.005,.72)	(.004,.74)	(.003,.75)	(.007,.73)	(.006,.72)

For the case  $N \geq 2$ , all procedures known to the authors for obtaining ordinary confidence intervals on the MOE depend on asymptotic distribution theory and hence are only approximate. For this reason, comparisons with the modified Bayesian intervals are meaningful only in the case of "large"  $n_i$ 's. With the exception of the last column giving the modified Bayesian limits, the figures in Table IV.4 appear in Harris (reference (7)). The modified Bayesian limits are seen to be close to the approximate ordinary limits and in fact are between Buehler's limit and Harris' limit for all but one set of the values of  $X_i$  and  $n_i$  used.

Table IV.4

Approximate Upper Limit of a Two-sided 80% Confidence Interval on  $MOE = P_1 P_2$  and  $MOE = P_1 P_2 P_3$  Where  $P_i$  is the Probability of Success of the  $i$ th Component

Sample Sizes $n_1, n_2$	Observed Values $x_1, x_2$	Beuhler's Limit	Madansky Likelihood Ratio	Madansky Linearized	Harris (Randomized)	Harris (Randomized)	Modified Bayesian ( $M=200$ )
100, 100	3, 5	.00412	.00433	.00164	.00486	.00416	.00444
100, 100	1, 4	.00188	.00182	.00097	.00235	.00184	.00292
100, 100	2, 2	.00168	.00167	.00091	.00211	.00170	.00193
150, 150	3, 3	.00128	.00133	.00074	.00153	.00128	.00142
Sample Sizes $n_1, n_2, n_3$		Observed Values $x_1, x_2, x_3$		Madansky Likelihood Ratio	Harris (Randomized)	Harris (Randomized)	Modified Bayesian ( $M=200$ )
100, 100, 100		1, 2, 1		.000019	.000040	.000027	.00002
100, 100, 100		2, 3, 5		.000133	.000186	.000145	.00014

If the  $n_i$ 's are large and the  $p_i$ 's are not close to zero then Madansky's likelihood ratio method is satisfactory. Table IV.5 contains one such example. Again, there is good agreement between the Modified Bayesian intervals and the ordinary intervals.

Table IV.5

Approximate Confidence Limits for MCE =  $\prod_{i=1}^N p_i$ , Where

$p_i$  = Probability of Success at ith Component

$$\prod_{i=1}^4 x_i/n_i = .156$$

$$x_1 = 34 \quad n_1 = 87$$

$$x_2 = 27 \quad n_2 = 33$$

$$x_3 = 23 \quad n_3 = 45$$

$$x_4 = 21 \quad n_4 = 22$$

Confidence Level	Madansky's Likelihood Ratio	Modified Bayesian (M=100)	Modified Bayesian (M=400)
80%	.116, .203	.097, .209	.101, .204
90%	.106, .218	.086, .218	.083, .223
95%	.098, .231	.043, .260	.044, .234

Unfortunately, numerical computations indicate that the modified Bayesian procedure is also severely biased if the  $p_i$ 's are close to one and the number of terms in the product,  $N$ , is large. Thus, it is recommended that the product  $\prod_{i=1}^N (x_i + 1) / (n_i + 2)$ , be "close" to the product,  $\prod_{i=1}^N (x_i / n_i)$ , before the modified Bayesian procedure is applied.

## V. A NOTE ON THE EFFECT OF "LATE STARTERS" ON THE ESTIMATE OF CUMULATIVE DETECTION PROBABILITY

### Discussion

In conducting the research for estimating confidence intervals on cumulative detection probability (CDP) as a function of range (see Chapter III of this paper), it was discovered that the inclusion of "late starters" in the data base may produce a biased estimate of the desired CDP. This bias is due to the fact that, in some cases, the "late starters" have their own CDP curve which may be significantly different from the CDP curve for "non-late starters". The magnitude of the bias of an estimate of CDP calculated from exercise data may be unknown.

### Theory

Let  $f(R)$  denote the cumulative detection probability (CDP) at range  $R$  of targets whose starting ranges are beyond the (reasonable) limits of detection. It is assumed that we desire to estimate the function  $f(R)$  for all  $R \geq 0$ . Let  $f_S(R)$  denote the CDP at range  $R$  of targets whose starting range is  $S$ , where  $S$  is less than the limits of detection. For ease of presentation assume there are two groups of targets: first,  $N$  targets beyond the limits of detection and then  $N_1$  targets start at range  $S_1$ . See Figure V.1.

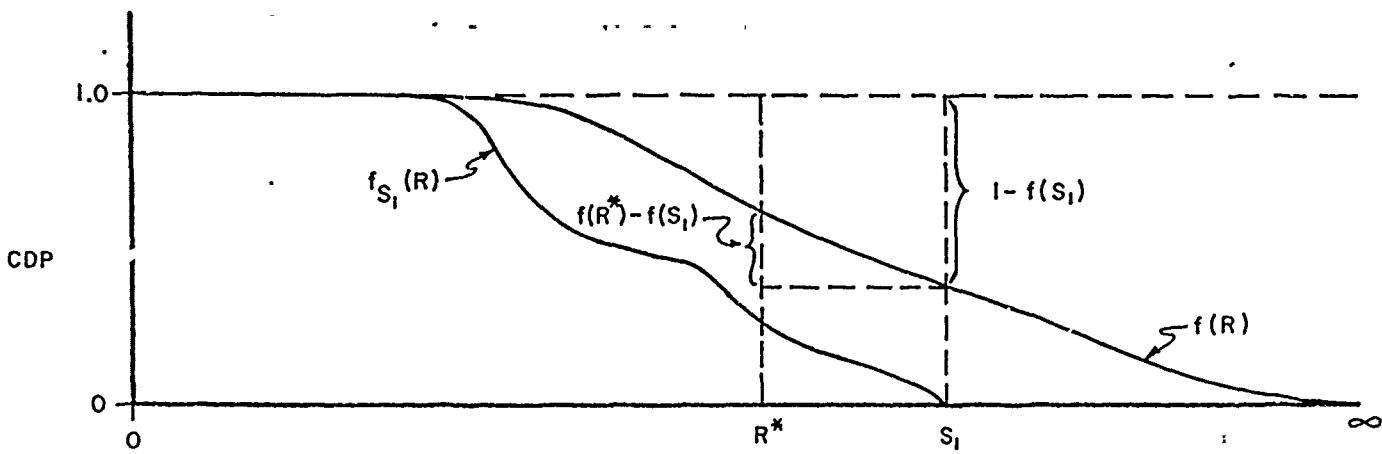


Figure V.1

Consider the range "bin"  $(R^*, S_1)$ . Using the standard technique in reference (1), the value of  $\hat{f}(R^*)$  can be estimated by  $\hat{f}(R^*)$ . The expected value of the estimate is

$$E(\hat{f}(R^*)) = 1 - \left( \frac{N(1-f(S_1))}{N} \right) \left( \frac{N(1-f(R^*)) + N_1(1-f_{S_1}(R^*))}{N(1-f(S_1)) + N_1} \right) \quad (V.1)$$

The right hand side of equation V.1 is not equal to  $f(R^*)$  for all possible values of the function  $f_{S_1}(R^*)$ , i.e., the estimation procedure is biased at  $R^*$  unless the function  $f_{S_1}(R)$  is such that  $E(\hat{f}(R^*)) = f(R^*)$ . Setting the right hand side of equation V.1 equal to  $f(R^*)$  and solving for  $f_{S_1}(R^*)$ , we find that

$$f_{S_1}(R^*) = (f(R^*) - f(S_1)) / (1-f(S_1)). \quad (V.2)$$

Equation V.2 is a necessary condition in order for the standard estimate,  $\hat{f}(R^*)$ , to be unbiased. If  $f_{S_1}(R^*) < (f(R^*) - f(S_1))/(1-f(S_1))$  then  $\hat{f}(R^*)$  is too small (on average) and if  $f_{S_1}(R^*) > (f(R^*) - f(S_1))/(1-f(S_1))$  then  $\hat{f}(R^*)$  is too large (on the average). The above remarks point out the fact that, if "late starters" have their own CDP curves then combining data from "late starters" with "non-late starters" to estimate the latter's CDP curve may produce biased results. The magnitude of the bias in a complicated exercise will be unknown.

To illustrate the above, two artificial but intuitive numerical examples follow. Assume that at range 6 miles the true CDP is 25% and at range 5 miles the true CDP is 30%. Assume that for late starters at range 6 miles, the probability is 25% that a target will be immediately detected. After the immediate detections, assume the CDP of the late starters follows the "true" CDP. See Figure V.2.

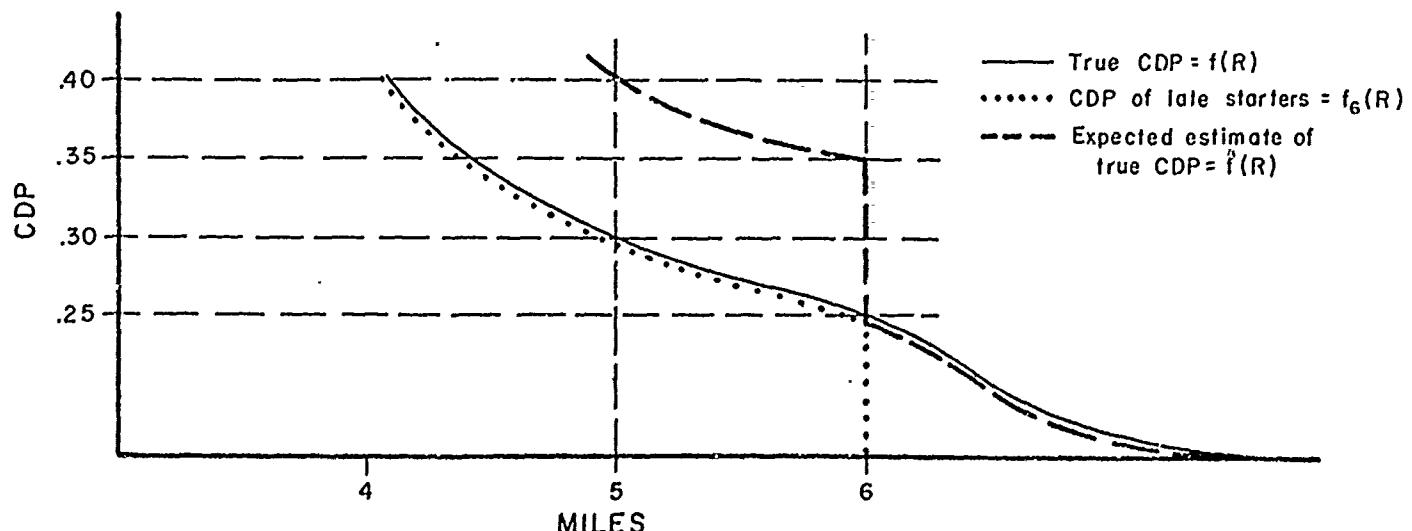


Figure V.2

In an exercise, assume there are 20 targets whose starting ranges are beyond the limits of detection and 20 targets which start at 6 miles. At a range "slightly" more than 6 miles, say  $6^+$ , the standard formulas will yield an estimate of CDP whose expected value is

$$E(\hat{f}(6^+)) = 1 - (15/20) = .25.$$

In this example, the estimation technique is unbiased for ranges greater than 6 miles. At 6 miles, 5 of the late starters will be detected immediately. In other words, we expect  $15 + 15 = 30$  nondetections out of  $15 + 20 = 35$  opportunities. The updated estimate of CDP will have the expected value

$$\begin{aligned} E(\hat{f}(6)) &= 1 - (15/20) (30/35) \\ &= .357. \end{aligned}$$

Thus, at range 6 miles, the estimation technique is biased by 10.7%. The "true" CDP is 25%, but on the average, the estimate of the true CDP is 35.7%. In the range "bin" from 6 to 5 miles, we expect  $(.05) (20) = 1$  detection from each group of targets. In other words, we expect 28 nondetections out of 30 opportunities. The updated estimate of CDP will have the expected value

$$\begin{aligned} E(\hat{f}(5)) &= 1 - (.643) (28/30) \\ &= 1 - .6 \\ &= .4. \end{aligned}$$

The procedure is biased by 10% at range 5 miles.

For the second example, assume that the probability is 1.0 that a target will be detected within one mile of starting range if its starting range is less than 8 miles. Further, assume that the probability of detection is uniform over this unit interval, and that there is no chance of detection of targets at a range greater than 8 miles. CDP curves for 3 starting ranges are sketched in Figure V.3.

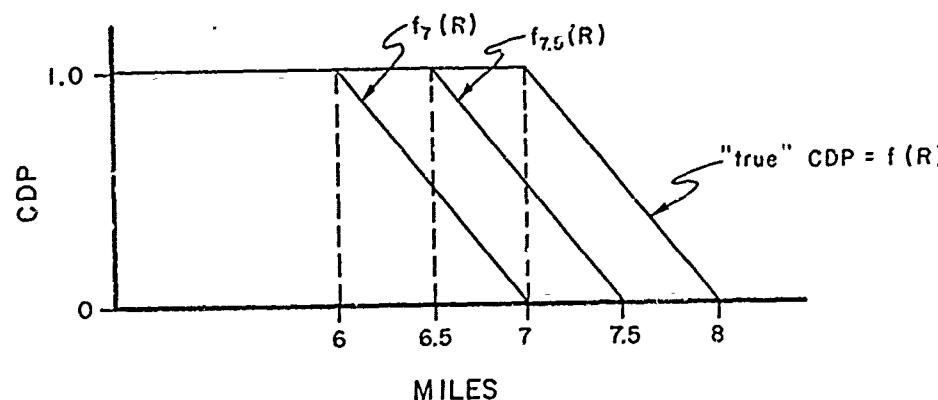


Figure V.3

Consider an exercise in which we have 10 targets with starting range greater than 8 miles, 5 late starters at 7.5 miles and 5 late starters at 7 miles. The expected CDP curve, as computed by the formulas in (reference (1)), appears in Figure V.4. For ranges less than 7.5 miles, the estimates are too small (on the average).

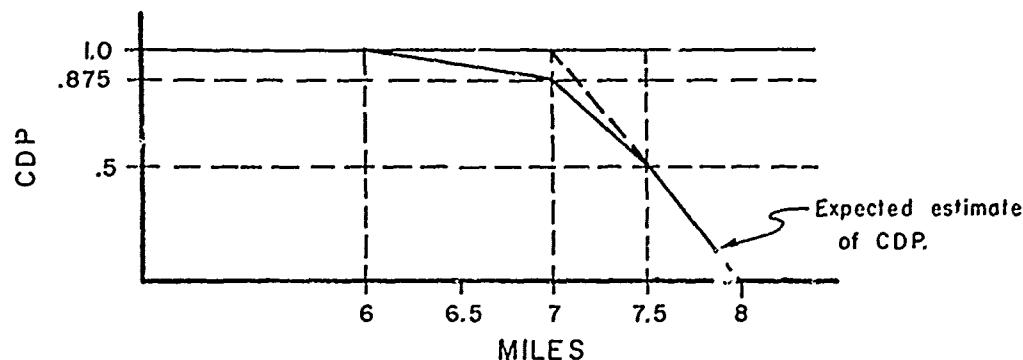


Figure V.4

#### Recommendation

An obvious solution to this potential bias due to "late-starters" is to eliminate all of them from the data base for calculating cumulative detection probability. However, this approach is not desirable since a significantly large portion of the data may consist of "late-starters", in which case the sample size would be drastically reduced. Thus, the analyst must use his discretion in eliminating some "late-starters" and retaining others.

The following guidelines for eliminating "late-starters" are recommended at this time:

- 1) All "late-starters" that are detected immediately (or just about immediately) after becoming a detection opportunity should be removed from the data base, since it is probable that they would have been detected at a longer range had they been opportunities at a longer range.
- 2) All "late-starters" that have a very short start range should be eliminated. A start range is considered to be very short if it is less than a large proportion of the detection ranges. (e.g., start range less than 50% of the detection ranges.)

Additional research relative to the "late-starter" effect on the distribution theory of cumulative detection probability may result in the formulation of different rules.

## REFERENCES

1. Submarine Analysis Notebook, by Commander Submarine Development Group TWO, CONFIDENTIAL.
2. Madansky, Albert, (1965). "Approximate Confidence Intervals for the Reliability of Series and Parallel Systems", Technometrics 7, pages 495-503.
3. Koopman, B. O. (1946). Search and Screening, OEG Report No. 56, Office of the Chief of Naval Operations, Navy Department, UNCLASSIFIED.
4. Hogg, R. V. and Craig, A. T. (1965). Introduction to Mathematical Statistics, Macmillan, New York.
5. Bossard, D. C. (1970). "Estimating Confidence Bands for Probability of Detection Curves", Daniel H. Wagner, Associates Memorandum Number 508/155.
6. Dyer A. R. (1968). "Estimates and Confidence Intervals for Curves of Cumulative Detection Probability", Daniel H. Wagner, Associates Memorandum Number 508/109.
7. Harris, Bernard (1971). "Hypothesis Testing and Confidence Intervals for Products and Quotients of Poisson Parameters with Applications to Reliability", J. Amer. Statist. Assoc. 66, pages 609-613.
8. Walsh, J. E. (1970). Approximate Intervals and Tests for Products and Ratios of Binomial Probabilities, Tech. Report No. 89, Dept. of Statistics, Southern Methodist University.
9. Buehler, R. J. (1957). "Confidence Intervals for the Product of Two Binomial Parameters". J. Amer. Statist. Assoc. 52, pages 482-493.
10. Springer, H. D. and Thompson, W. E. (1966). "Bayesian Confidence Limits for the Product of N Binomial Parameters", Biometrika 53, pages 611-613.
11. Mood, A. M. and Graybill, F. A. (1963). Introduction to the Theory of Statistics, McGraw Hill, New York.
12. Wilks, S. S. (1962). Mathematical Statistics, Wiley, New York.
13. Myhre, J. M. and Saunders, S. C. (1968). "Comparison of Two Methods of Obtaining Approximate Confidence Intervals for System Reliability", Technometrics 10, pages 37-49.